

Apunts sobre proves de rendiment

Característiques, avantatges i limitacions

Francesc Martínez-Olmo
2021



UNIVERSITAT DE
BARCELONA

Aquesta obra està subjecta a una llicència
de [Reconeixement 4.0 Internacional de Creative Commons](#)



Taula de contingut

1	<i>Proves d'assaig</i>	3
2	<i>Proves objectives</i>	4
2.1	Etapas per a l'elaboració d'una prova objectiva	5
2.2	Tipus de preguntes objectives	5
3	<i>Taxonomia d'objectius cognoscitius</i>	10
4	<i>Proves alternatives de rendiment</i>	13
4.1	El portafolis com a paradigma de l'avaluació alternativa	14
5	<i>Bibliografia per saber-ne més</i>	17
6	<i>Referències</i>	18

Aquest document ha estat corregit pels Serveis Lingüístics en el marc de la convocatòria de correcció de materials docents del Pla de llengües de la Universitat de Barcelona.

1 Proves d'assaig

El rendiment educatiu s'ha avaluat tradicionalment a partir de proves com els assajos —exàmens tradicionals on les preguntes s'han de contestar de manera relativament extensa—, les proves objectives i l'observació directa de l'estudiant. En aquest apartat es tracten alguns dels fonaments de les proves d'assaig.

El concepte d'assaig queda reflectit en la definició que aporta Merino (1992, p. 221, citat a Mateo i Martínez-Olmo, 2008b):

Entendemos por ensayo una composición escrita de extensión variable —generalmente breve— cuya finalidad es la comunicación de un mensaje o idea principal, en conexión con un tema de estudio, en torno al cual debe seleccionarse el estilo más adecuado a la idea principal y a su desarrollo y utilizando la reflexión personal y el dominio de la materia de estudio para apoyar la comunicación sobre fundamentos sólidos...

Les preguntes d'assaig es classifiquen en funció de l'amplitud de desenvolupament que sol·liciten; d'aquesta manera es diferencia entre les preguntes de resposta àmplia i les de resposta restringida. A més a més, els treballs monogràfics que se solen demanar en certes avaluacions es consideren també assajos de resposta àmplia. La frontera que distingeix la resposta restringida de l'àmplia és relativa i difusa, però es pot entendre que una resposta d'una extensió aproximada entre una frase i un o dos paràgrafs correspondria a una resposta restringida i, quan l'extensió aproximada és de més d'un o dos paràgrafs, aleshores es tractaria d'una resposta àmplia.

Per construir una prova d'assaig s'aconsella seguir les indicacions següents:

1. Restringir l'ús d'aquest tipus de prova a aquells aprenentatges que no es puguin avaluar satisfactòriament amb proves objectives.
2. Formular preguntes que impliquin l'aprenentatge especificat als objectius del procés d'ensenyament i aprenentatge.
3. Plantejar cada pregunta de manera que quedi clarament especificada la tasca que ha de fer l'estudiant. Si cal, subdividir el tema de l'assaig en diverses parts i indicar clarament els punts que es poden aconseguir a cada apartat.
4. Indicar un límit aproximat de temps o d'espai per a cada pregunta o per a l'assaig sencer.
5. Indicar els factors que poden influir en la puntuació —llegibilitat, faltes d'ortografia, ús de vocabulari professional, etc.

Per puntuar les respostes de les proves d'assaig, s'aconsella:

1. Preparar anticipadament un esborrany de les respostes correctes, que serveixi de patró o model.
2. Especificar una puntuació per a cadascuna de les parts de la resposta.

3. Puntuar pregunta per pregunta de tots els estudiants —puntuar la primera pregunta de tots els estudiants, després la segona pregunta de tots els estudiants, etc.
4. Puntuar de manera anònima.
5. Puntuar l'assaig per part de més d'un avaluador, si l'avaluació és una condició necessària per prendre decisions importants.

L'elaboració d'un assaig de resposta àmplia no és una tasca evident per a moltes persones. Alguns criteris que orienten l'elaboració i l'avaluació d'un assaig són:

1. Idea principal. La constitueix tota hipòtesi, idea nuclear, problema, finalitat o propòsit que es desitja estudiar i transmetre amb l'elaboració de l'assaig. Respon a què, per què i per a què s'ha fet el treball. Fa saber des de quina perspectiva i de quina manera es tracta la qüestió, tema o problema plantejat.
2. Contingut. Té a veure amb el conjunt d'informació —idees, fets, dades, etc.— que s'inclou a l'assaig, de la qual és important la quantitat i la qualitat. La selecció del contingut ha de ser pertinent i oportuna, i ha d'estar relacionada directament amb l'assaig.
3. Estructura. L'ordre i la distribució de les parts d'un assaig han de tenir relació amb l'argumentació que es presenta.
4. Raonament. Es refereix a la manera lògica, coherent i ben fonamentada d'exposar i argumentar una qüestió a partir d'una idea principal per assolir una conclusió.
5. Expressió. Es refereix a la manera de transmetre les idees. Implica un estil fluid, concís i clar.

El valor pedagògic de les proves d'assaig incrementa en gran mesura si la persona avaluada rep, en un termini de temps relativament breu, comentaris específics i argumentats sobre les respostes. Els criteris que s'acaben d'exposar orienten també sobre el tipus de comentaris que es poden retornar.

Algunes de les crítiques que s'han fet a les proves d'assaig se centren en la falta d'uniformitat en el moment de puntuar les respostes (fiabilitat), fet que mostra la varietat existent de criteris. També s'ha qüestionat la seva validesa, entesa com a representativitat, sobretot quan es fan poques preguntes en el conjunt de la prova d'assaig.

2 Proves objectives

Una prova objectiva és un tipus d'examen escrit estructurat que té les característiques principals següents: presenta un nombre elevat de preguntes, té una estructura tancada de resposta i les respostes acceptables són unívokes i exactes, cosa que permet que es pugui qualificar objectivament.

2.1 Etapes per a l'elaboració d'una prova objectiva

L'elaboració d'un test o prova objectiva és un procés complex en què intervenen diversos factors. En funció del coneixement previ sobre la teoria que doni suport a allò que es vol avaluar, dels objectius del test i fins i tot dels recursos humans i econòmics disponibles, les fases per construir un test poden variar d'un cas a un altre, però en general es poden resumir com es mostra a la taula següent.

Taula 1

Fases en la construcció d'una prova objectiva

Fase	Tasca	Concreció
Planificació	<ul style="list-style-type: none"> • Determinar la finalitat de la prova 	<ul style="list-style-type: none"> • Diagnòstica, formativa, sumativa
	<ul style="list-style-type: none"> • Delimitar els objectius i continguts d'aprenentatge 	<ul style="list-style-type: none"> • Especificació, tan concreta com sigui possible, dels objectius objecte d'avaluació
Elaboració del protocol	<ul style="list-style-type: none"> • Determinar la importància relativa dels objectius i continguts en el conjunt de la prova 	<ul style="list-style-type: none"> • Construcció de la taula d'especificació • Determinació de la quantitat de preguntes
	<ul style="list-style-type: none"> • Redactar les preguntes 	<ul style="list-style-type: none"> • Selecció del tipus de preguntes • Organització de la seqüència de les preguntes
	<ul style="list-style-type: none"> • Elaborar el «quadern» (protocol) de la prova, amb les instruccions d'aplicació i correcció 	
Anàlisi de les preguntes	<ul style="list-style-type: none"> • Dur a terme l'anàlisi dels ítems • Suggerir millores en el redactat de la prova o de les preguntes 	<ul style="list-style-type: none"> • Càlcul dels índexs de facilitat, discriminació i homogeneïtat, i anàlisi de distractors

2.2 Tipus de preguntes objectives

Hi ha diversos tipus de preguntes que formen una prova objectiva. Alguns dels més utilitzats són:

1. Ítem de llacuna o resposta breu. Es caracteritza per establir una proposició incompleta o una pregunta que s'ha de completar o respondre amb una paraula, frase, símbol, número o resposta breu.
2. Ítem de resposta alterna —sí o no, vertader o fals. La resposta es limita a dues opcions que qualifiquen un enunciat i s'ha d'escollir la correcta.

3. Ítem d'ordenació. S'ofereix una llista d'elements que s'han d'ordenar d'acord amb algun criteri indicat a l'ítem —cronològic, lògic, operatiu, espacial, etc.
4. Ítem d'aparellament o correspondència. Requereix establir les relacions entre els elements de dos grups.
5. Ítem d'elecció múltiple. És una pregunta breu, seguida de diverses alternatives de resposta, de les quals només una és la correcta o la millor —també es pot donar el cas que hi hagi més d'una elecció correcta i s'hagin de seleccionar totes—; la versió més recomanable és la de resposta única correcta, amb dos distractors,¹ que és com s'anomenen les alternatives falses.

Disposem de diversos manuals que expliquen les normes de redacció per a cadascun dels tipus d'ítems objectius (Blanco, 1994; Cohen et al., 2007; del Rincón et al., 1995; Espín i Rodríguez Lajo, 1996; Renom, 1992). A continuació es mostren les característiques principals de cada tipus d'ítem:

Taula 2

Ítem de llacuna o resposta breu

Utilitat
<ul style="list-style-type: none"> Es pot utilitzar per avaluar objectius que impliquin memoritzar dades, símbols, fórmules, etc.
<p>Criteris per a l'elaboració</p> <ul style="list-style-type: none"> Es recomana utilitzar preferentment la modalitat de resposta breu, a partir d'una pregunta específica. La pregunta ha de conduir a una única resposta. No s'ha de formular amb les expressions i termes que figuren literalment en els documents utilitzats per a l'aprenentatge. Per facilitar la correcció, convé situar totes les línies de resposta a la dreta del full. La pregunta no ha de donar indicis de resposta. En cas que sigui un text mutilat: <ul style="list-style-type: none"> L'espai que ocupen totes les llacunes ha de tenir la mateixa llargada. Les llacunes han de ser prou grans perquè hi càpiga la resposta. La llacuna no ha d'anar al principi de la frase.
<p>Avantatges</p> <ul style="list-style-type: none"> És fàcil d'elaborar. S'elimina força la possibilitat de respondre correctament a l'atzar. Es pot aplicar també a mapes, dibuixos, gràfiques, etc. Es corregeix amb facilitat. Es redueix molt l'arbitrarietat en la correcció.

¹ Per elaborar alternatives falses o distractors, es pot utilitzar una estratègia que consisteix a trobar les respostes a què s'arribaria cometent un error en els diversos passos del procés per resoldre el problema plantejat a l'ítem. En aquest cas, quan la persona que respon escull una alternativa falsa, es pot saber la naturalesa de l'error que comet i, consegüentment, es pot oferir una orientació específica. D'aquesta manera es dona un caràcter més formatiu a l'avaluació.

Inconvenients
<ul style="list-style-type: none"> No facilita l'avaluació d'aprenentatges complexos.
Exemples
<p>a) Quin és el pH d'una dissolució neutra? _____</p> <p>b) Quevedo va néixer l'any _____</p> <p>c) El sonet està format per dos _____ i dos _____</p>

Taula 3

Ítem de resposta alterna

Utilitat
<ul style="list-style-type: none"> S'utilitza per identificar relacions causa-efecte i per distingir fets i opinions.
Criteris per a l'elaboració
<ul style="list-style-type: none"> L'enunciat ha de ser tan categòric com sigui possible, de manera que només sigui estrictament acceptable una de les alternatives. S'han de redactar més enunciats falsos que vertaders, ja que discriminen més. Convé que l'enunciat sigui breu, senzill i fàcil d'entendre. S'ha d'evitar que la resposta correcta es pugui donar utilitzant simplement el sentit comú. S'han d'evitar les dobles negacions en les proposicions. Els enunciats haurien de tenir una extensió similar. El redactat no ha de ser un fragment literal del document que s'ha utilitzat en l'aprenentatge. Cada ítem ha de plantejar una sola pregunta.
Avantatges
<ul style="list-style-type: none"> És fàcil de construir. Es respon amb facilitat. Es pot utilitzar en tots els nivells educatius.
Inconvenients
<ul style="list-style-type: none"> L'atzar hi intervé molt. Només permet recollir informació de petits àmbits del coneixement. Té un escàs valor diagnòstic. Les persones avaluades tenen més tendència a seleccionar respostes negatives que positives.
Exemple
<ul style="list-style-type: none"> L'aigua pura, a pressió atmosfèrica, bull a 1.000 °C: V o F

Taula 4

Ítem d'ordenació

Utilitat
<ul style="list-style-type: none"> • És particularment útil per a continguts procedimentals.
Criteris per a l'elaboració <ul style="list-style-type: none"> • S'aconsella escriure primer les frases ordenadament i després distribuir-les a l'atzar. • Convé contrastar l'ítem amb algun col·lega abans d'utilitzar-lo en una avaluació, per verificar el grau d'acord en l'ordenació. • S'han d'evitar termes que ajudin a respondre, com ara «després», «a continuació», «finalment», etc.
Avantatges <ul style="list-style-type: none"> • Permet comprovar l'ordre mental que té la persona que respon.
Inconvenients <ul style="list-style-type: none"> • De vegades no són fàcils de puntuar —com es valora que hi hagi un element desordenat? I si n'hi ha més d'un? • Generalment són difícils de construir.
Exemples <ul style="list-style-type: none"> • Ordena de més a menys duresa els minerals següents: <ul style="list-style-type: none"> ○ calcita, ortosa, talc, diamant, topazi, guix

Taula 5

Ítem d'aparellament o correspondència

Utilitat									
<ul style="list-style-type: none">• Permet relacionar substàncies i propietats, causes i efectes, termes i definicions, successos i dates, aparells i funcions, etc.									
Criteris per a l'elaboració									
<ul style="list-style-type: none">• A la segona columna hi ha d'haver més elements que a la primera.• Les expressions han de ser homogènies en cada columna.• La relació entre les expressions ha de ser unívoca.									
Avantatges									
<ul style="list-style-type: none">• Permet comprovar força coneixements en poc temps.• Permet utilitzar dibuixos, gràfiques, etc.									
Inconvenients									
<ul style="list-style-type: none">• No és fàcil de construir.• No és adequat per a aprenentatges complexos.									
Exemple									
<ul style="list-style-type: none">• Associa cada país amb un continent —quedarà un continent sense relacionar:<table><tr><td>Holanda</td><td>Europa</td><td>Àfrica</td></tr><tr><td>Egipte</td><td>Àsia</td><td></td></tr><tr><td>Colòmbia</td><td>Amèrica</td><td></td></tr></table>	Holanda	Europa	Àfrica	Egipte	Àsia		Colòmbia	Amèrica	
Holanda	Europa	Àfrica							
Egipte	Àsia								
Colòmbia	Amèrica								

Taula 6**Ítem d'elecció múltiple**

Utilitat <ul style="list-style-type: none"> • Permet avaluar diversos nivells d'aprenentatge —memòria, comprensió, anàlisi, síntesi, avaluació.
Criteris per a l'elaboració <ul style="list-style-type: none"> • Pel que fa a la base de l'ítem: <ul style="list-style-type: none"> ○ La redacció ha de ser clara i breu. ○ S'ha de referir a un sol contingut. ○ No s'ha de redactar literalment tal com apareix en el material d'aprenentatge. ○ Convé redactar-la com una pregunta, no com una frase inacabada. ○ No s'ha d'expressar en mode negatiu. ○ No ha de contenir informació aliena al que es pregunta només per allargar l'ítem. ○ Una pregunta no ha d'ajudar a respondre una altra pregunta. • Pel que fa a les opcions de resposta: <ul style="list-style-type: none"> ○ Tots els ítems han de tenir la mateixa quantitat d'opcions de resposta. ○ No s'han de repetir paraules. ○ Totes les opcions de resposta d'un mateix ítem han de mantenir paral·lelisme gramatical i coherència amb el peu de l'ítem. ○ Només una opció hauria de ser correcta. ○ Els distractors —opcions incorrectes— han de ser creïbles. ○ Totes han de tenir la mateixa extensió aproximada. ○ Les respostes correctes s'han de distribuir a l'atzar. ○ Evitar les opcions «totes les anteriors»², «cap de les anteriors».³
Avantatges <ul style="list-style-type: none"> • Informa sobre aprenentatges complexos si estan ben redactades. • No és tan fàcil de respondre bé a l'atzar. • Permet puntuar objectivament.
Desavantatges <ul style="list-style-type: none"> • És difícil de construir. • És complicat trobar bons distractors. • No facilita l'avaluació de la creativitat.

² Quan s'inclou l'opció «totes les anteriors» incrementa la probabilitat d'encertar a l'atzar, per dos motius: a) si aquesta opció és la correcta, és més fàcil encertar pel fet que la persona que respon detecti que dues de les altres opcions són correctes, a més a més hi ha el problema que s'incompleix la consigna que digui que només una de les opcions és correcta; b) en cas que aquesta opció sigui incorrecta, també és més fàcil encertar només pel fet que la persona que respon detecti que una de les opcions és falsa.

³ Quan s'inclou l'opció «cap de les anteriors», per una banda, s'incrementa la probabilitat d'encertar a l'atzar si la persona que respon detecta que una de les opcions podria ser correcta; i per l'altra, si aquesta és l'opció correcta, no sabem si la persona que respon sap realment quina seria la resposta correcta a la pregunta plantejada.

Exemple

En quina unitat del sistema internacional es mesura la força?

- a) Kilopond
- b) Newtòmetre
- c) Newton
- d) Kilonewton

En les preguntes de resposta alterna i en les d'elecció múltiple, per tal d'evitar que es responguin confiant en la sort que s'encertarà la resposta correcta, s'aplica el que s'anomena correcció de l'atzar, de manera que les omissions —el fet de no respondre— no descompten, però cada error resta d'acord amb la fórmula següent:

$$\text{Puntuació}_{\text{error}} = \frac{\text{Valor de l'ítem}}{\text{Núm. d'alternatives} - 1}$$

Per exemple, en una pregunta de tipus test amb tres opcions (*a*, *b* i *c*), suposant que la pregunta val 1 punt, un error ha de descomptar ½ punt, és a dir, 0,5 punts.

3 Taxonomia d'objectius cognoscitius

Una classificació o taxonomia d'objectius serveix tant per elaborar com per validar una prova d'avaluació. L'interès recau a diferenciar, amb un criteri determinat, diversos tipus d'objectius, de manera que s'aconsegueixen proves d'avaluació que tenen en consideració un ventall més ampli d'aprenentatges.

La taxonomia de Bloom, publicada per primer cop l'any 1956, va tenir un impacte extraordinari, perquè va sistematitzar d'una manera senzilla tots els nivells d'aprenentatge cognoscitiu. Fins a aquell moment, l'interès de l'ensenyament i l'avaluació se centrava sobretot en els nivells més baixos d'aprenentatge —memòria i comprensió—, però aquesta taxonomia va ajudar a equilibrar-ho fent posar atenció als nivells superiors dels processos cognitius —vegeu taula següent.

Les taxonomies d'objectius també són útils per planificar i comprovar la validesa de contingut de les proves de rendiment. A la fase inicial d'elaboració d'una prova es decideix sobre quins continguts i quins nivells d'aprenentatge es desitja obtenir informació, i a la fase de validació es comprova si la quantitat d'ítems que mesuren cadascun dels objectius i dels nivells es correspon amb allò que es necessita avaluar. Per facilitar aquestes tasques es construeix una taula d'especificació com la de l'exemple següent.

Taula 7

Exemple d'una taula d'especificació d'una prova de rendiment sobre informàtica

Contingut	Objectius			Ponderació de continguts
	Memòria	Comprensió	Aplicació	
Adaptació del maquinari	1	3	2, 4, 8, 6	6 (12 %)
Memòries USB	7, 5		9, 10, 11, 12, 13	7 (14 %)
Organització de l'escriptori	14		15, 16, 17, 18	5 (10 %)
Execució del programari	19, 20, 21	22, 23	24, 25, 26, 27, 28, 29, 30, 31	13 (32 %)
Configuració del sistema operatiu	32, 33, 34, 35, 36	37	38, 39, 40, 41, 42, 43, 44, 45, 46, 47	16 (32 %)
Gestió d'errors	48		49, 50	3 (6 %)
Ponderació d'objectius	13 (26 %)	4 (8 %)	33 (66 %)	50 ítems (100 %)

La taxonomia de Bloom ha estat criticada perquè es formula en termes conductuals —aprenentatge conductista—; la separació entre els aspectes cognitius, actitudinals i procedimentals no és real i, per tant, pot orientar incorrectament l'ensenyament. Així i tot, se li reconeix el seu valor perquè aporta un marc organitzat en què els objectius es classifiquen amb alguna precisió.

A continuació es mostra una adaptació de la taxonomia d'objectius cognoscitius de Bloom al procés reflexiu:

Figura 1

Taxonomia d'objectius sobre la reflexió

- | | |
|----------------|--|
| 1. Memòria: | Què s'ha fet? |
| 2. Comprensió: | Què és important del que s'ha fet? S'han aconseguit els objectius? |
| 3. Aplicació: | Quan s'ha dut a terme abans? On el podria usar de nou? |
| 4. Anàlisi: | S'observen patrons o relacions en el que s'ha dut a terme? |
| 5. Avaluació: | Com de bé s'ha dut a terme? Què ha funcionat? Què ha de millorar? |
| 6. Creació: | Què podria fer a continuació? Quin és el meu pla o disseny? |

Nota. Adaptat de Pappas (2010).

Categorías de la taxonomía de Bloom	Ejemplos de objetivos	Ejemplos de términos
<p>1. Saber. El saber se define como el acto de recordar el material previamente aprendido. Hacerlo así presupone recordar una rica gama de materiales que van desde hechos concretos hasta teorías completas pero, en todo caso, lo que se requiere es traer a colación la información apropiada. El saber representa el nivel más bajo de los productos del aprendizaje en el dominio cognoscitivo.</p> <p>2. Comprensión. La comprensión se define como la capacidad para captar el significado del material. Dicha comprensión puede demostrarse al traducir un material de una forma a otra (palabras a números), al interpretar el material (resumirlo o resumirlo) y al estimar cuáles serán las futuras tendencias (predicción de consecuencias o efectos). Estos productos del aprendizaje van un paso más allá de la simple recordación del material y representan al nivel más bajo de la comprensión.</p> <p>3. Aplicación. La aplicación se refiere a la capacidad de usar el material aprendido en situaciones nuevas y concretas. Hacerlo así puede muy bien incluir la aplicación de elementos tales como reglas, métodos, conceptos, principios, leyes y teorías. Los productos del aprendizaje en esta área requieren un nivel más alto de comprensión que el caso anterior.</p> <p>4. Análisis. El análisis se refiere a la capacidad de subdividir el material dado en las partes que lo componen de manera que pueda comprenderse la estructura de su organización. Hacerlo así puede incluir la identificación de las partes, el análisis de las relaciones entre las partes y el reconocimiento de los principios de organización involucrados. Los productos del aprendizaje representan aquí un nivel intelectual más elevado que en la comprensión y la aplicación, dado que requieren la comprensión tanto del contenido como de la forma estructural del material.</p> <p>5. Síntesis. La síntesis se ocupa de la capacidad de juntar las partes de manera de formar un nuevo todo para lo cual puede requerirse la producción de una comunicación única (tema o discurso), un plan de operaciones (proposición de investigaciones) o un conjunto de relaciones abstractas (plan para clasificar información). Los productos del aprendizaje en esta área recalcan los comportamientos creativos, recayendo su mayor énfasis en la formulación de nuevos patrones o estructuras.</p> <p>6. Evaluación. A la evaluación le atañe la capacidad de juzgar el valor del material (aseveración: novela, poema, informe de investigaciones) para un propósito dado. Los juicios han de fundamentarse en criterios definidos y pueden ser internos (de organización) o externos (pertinencia con respecto a la finalidad) y al estudiante pueden dársele los criterios o permitirle que los determine por su cuenta. Los productos del aprendizaje en esta área son los más elevados en la jerarquía cognoscitiva porque contienen elementos de todas las demás categorías a los que se añaden juicios conscientes de valores que se apoyan en criterios claramente definidos.</p>	<p>Conoce términos comunes Conoce hechos específicos Conoce métodos y procedimientos Conoce conceptos básicos Conoce principios Comprende hechos y principios Interpreta material verbal Interpreta gráficas y tablas Traduce el material verbal a fórmulas matemáticas Estima las futuras consecuencias que los datos implican Justifica métodos y procedimientos Aplica conceptos y principios a situaciones nuevas Aplica leyes y teorías a situaciones prácticas Resuelve problemas matemáticos Construye gráficas y tablas Demuestra el uso correcto de un método o de un procedimiento Reconoce hipótesis no expresadas Reconoce las falacias en el razonamiento lógico Distingue entre hechos e inferencias Evalúa la pertinencia de los datos Analiza la estructura organizada de una obra (arte, música, literatura) Escribe un tema bien estructurado Presenta un discurso bien estructurado Escribe una narración breve creativa (o un poema o una composición musical) Propone un plan para un experimento Integra el aprendizaje de áreas diferentes en un plan para resolver un problema Formula un nuevo plan para clasificar objetos (o eventos o ideas) Juega sobre la consistencia lógica del material escrito Emite un juicio sobre el grado de adecuación con que los datos apoyan las conclusiones Juega el valor de una obra (arte, música, literatura), usando criterios internos Conceptúa el valor de una obra (arte, música, literatura), usando normas externas de excelencia</p>	<p>Define, describe, identifica, rotula, forma listas, casa, nombra, esboza, reproduce, selecciona, expresa, assera Convierte, doblende, distingue, estima, explica, prolonga, generaliza, ejemplifica, infiere, parafrasea, predice, vuelve a redactar, resume Cambia, calcula, transforma, demuestra, descubre, manipula, modifica, monaja, predice, prepara, produce, relaciona, demuestra, resuelve, utiliza Subdivide, hace diagramas, diferencia, discrimina, distingue, identifica, ilustra, infiere, esboza, destaca, menusa Forma categorías, combiaa, compila, compone, crea, idea, diseña, explica, genera, modifica, organiza, planea, vuelve a ordenar, reconstruye, relaciona, reorganiza, revisa, vuelve a redactar, resume, planifica, escribe Tasa, compara, llega a conclusiones, contrasta, critica, describe, discrimina, explica, justifica, interpreta, relaciona, resume, apoya</p>

Figura 2
Taxonomía d'objectius i exemples

4 Proves alternatives de rendiment

A finals del segle xx, l'avaluació dels aprenentatges es va veure immersa en un conjunt de canvis que en van afectar tant la concepció com la metodologia. A l'apartat anterior s'ha tractat un dels instruments clàssics d'avaluació del rendiment, el test o prova objectiva, que juntament amb les proves d'assaig —orals o escrites—, en què l'alumnat respon de forma oberta i àmplia a una o diverses preguntes o temes, es constituïen com els únics instruments necessaris per al tipus d'avaluació requerida pel sistema educatiu. Actualment, en canvi, el sistema educatiu s'interessa per altres aspectes:

- a) comprovar si els estudiants adquireixen els coneixements, les competències i les habilitats que garanteixin una bona inserció sociolaboral o en nivells superiors de formació;
- b) avaluar continguts procedimentals i actitudinals, específics i transversals, per exemple, el pensament crític, la resolució de problemes, la comunicació escrita i oral o la competència tecnològica;
- c) millorar la institució educativa en el seu conjunt; per aquest motiu la informació que aporta l'avaluació és clau, encara que això requereixi canvis organitzatius i un esforç per part del professorat.

Les principals diferències que suposa aquest canvi de concepció en l'avaluació es resumeixen, seguint la proposta de Wolf i Reardon (1996), en la taula següent:

Taula 8

Diferències entre avaluació tradicional i alternativa

Característica	Avaluació tradicional	Avaluació alternativa
Mostres	Proves objectives	Experiments, debats, portafolis, productes dels estudiants
Base del judici	Informació objectiva i interpretació de puntuacions	Observació, subjectivitat, judici professional
Focus	Puntuació de l'estudiant i comparació amb altres puntuacions	Execució personalitzada de l'estudiant
Resultat	Una puntuació	Una història avaluativa respecte de l'individu o del grup
Interpretació	Generalitzable	Idiosincràtica i reflexiva
Conseqüència	Inhibeix l'acció curricular	Facilita l'acció i el canvi curricular
Poder	El professorat o una altra força externa té tot el poder	L'estudiant participa en la seva pròpia avaluació
Temps dedicat	Relativament breu	Requereix força temps

L'avaluació alternativa es considera habitualment com a sinònima d'autèntica (*authentic*) i d'execució (*performance*) (Pett, 1990). Aquest tipus d'avaluació emfatitza l'ús de mètodes que faciliten l'observació directa de la feina i de les habilitats de l'alumnat (Mateo, 2000). Alguns especialistes diferencien entre l'avaluació autèntica i la d'execució, argumentant que la primera requereix que l'alumnat executi alguna tasca que pertanyi a la vida real o que s'hi assembli molt, mentre que la segona només requereix l'execució d'algun exercici que, encara que potser mai s'ha de fer a la vida real, posi en pràctica les habilitats apreses (Meyer, 1992).

L'avaluació autèntica o d'execució es pot implementar formalment o informal. En el primer cas l'alumnat sap que està sent avaluat, mentre que en el segon cas l'alumnat no se n'adona. De fet, en molts casos l'alumnat és avaluat informalment, per exemple, quan és observat en la seva interacció amb els companys.

En el nostre cas diferenciem l'avaluació alternativa de l'avaluació autèntica i d'execució, considerant, d'acord amb Mateo (2000), com a «alternativa» qualsevol avaluació que apliqui instruments de mesura diferents de les proves d'elecció múltiple o del clàssic examen de preguntes obertes.

L'avaluació alternativa adquireix encara més sentit si descentrem el focus d'atenció sobre l'avaluació sumativa per passar a pensar també en l'avaluació diagnòstica i formativa. Els instruments alternatius d'avaluació del rendiment admeten fàcilment múltiples adaptacions segons el context específic d'aplicació, fet que els permet tenir més validesa, en funció dels objectius d'aprenentatge proposats, encara que redueix la possibilitat d'estandardització.

La utilitat —o validesa en termes de la teoria clàssica dels tests— d'aquest tipus d'avaluació alternativa s'argumenta sobre la base de la similitud dels instruments alternatius amb les tasques o els problemes que s'han de solucionar en la vida real. Si l'objectiu de l'avaluació és simplement comprovar els resultats, l'avaluació tradicional és útil, però si a més es pretén millorar la pràctica, l'avaluació alternativa ofereix millors orientacions i, per tant, és encara més útil. Pel que fa a la precisió dels resultats obtinguts, deixant de banda els errors que sempre es produeixen en qualsevol mesurament de fenòmens humans, l'avaluació alternativa sempre facilita els productes observables directament, de manera que qualsevol persona pot comprovar l'ajust de les interpretacions fetes. Així mateix, l'avaluació alternativa es caracteritza per l'ús de múltiples fonts d'informació, amb la qual cosa també s'aconsegueix més validesa i fiabilitat.

4.1 El portafolis com a paradigma de l'avaluació alternativa

El portafolis no és un concepte nou. Cap als anys 80 del segle xx es va renovar l'interès pels portafolis a causa de la percepció que aquest instrument millorava l'avaluació i motivava l'alumnat en el seu propi procés d'aprenentatge (Arter, 1995).

La definició de portafolis varia segons l'autor que es consulti, però hi ha un cert consens a considerar-lo com una col·lecció de materials amb la intenció d'explicar i reflexionar sobre el rendiment o l'aprenentatge realitzat.

El portafolis comporta un concepte d'educació alternatiu perquè dona veu als que normalment no poden mostrar el seu aprenentatge a través de les clàssiques avaluacions i, en cert sentit, retorna el poder als exclosos del sistema educatiu (Sunstein i Lovell, 2000).

L'interès d'utilitzar el portafolis com a instrument d'avaluació es deu al fet que proporciona múltiples mostres de la feina feta, la qual cosa permet obtenir:

- Una visió més àmplia i profunda del que se sap i del que es pot fer.
- Una base per a l'avaluació autèntica.
- Un mitjà per aportar evidències de les competències adquirides.
- Un complement per a la informació obtinguda a través d'avaluacions tradicionals (exàmens).
- Un reflex de les actituds i els valors, a més del coneixement i les habilitats, de la persona que elabora el portafolis.
- Una altra manera de comunicar el progrés a altres agents —direcció, inspecció, família, ocupadors, etc.
- Una eina per al propi desenvolupament professional.
- Una eina per a la millora, ja que en facilitar la reflexió sobre el que s'ha après, s'amplien les possibilitats de canviar.

Algunes de les utilitats que s'han donat als portafolis són:

- Comprovar el desenvolupament realitzat.
- Reflexionar i aprendre sobre el desenvolupament realitzat, en autoavaluacions.
- Planificar un programa, utilitzant el portafolis com a diagnòstic.
- Aportar evidències i exemplificar les habilitats millor apreses per sol·licitar un treball o per ingressar en una universitat.

Normalment en un portafolis no s'inclouen «tots» els treballs realitzats, sinó que es fa una selecció en funció de l'objectiu fonamental del portafolis, que pot variar des de promoure l'autoavaluació, fins a avaluar un programa. L'estructura bàsica del portafolis gira al voltant de tres eixos: *a)* taula de continguts o mapa del portafolis; *b)* una col·lecció de material seleccionat i il·lustratiu de la pràctica realitzada, i *c)* una reflexió comentant les evidències que aporta el material del portafolis. A tall d'exemple, l'estructura d'un portafolis pot tenir els apartats següents:

- Portada i taula de continguts

- Marc o context d'aprenentatge
- Objectius d'aprenentatge, responsabilitats i compromisos adquirits amb referència als objectius
- Metodologies utilitzades
- Exemples de materials realitzats
- Evidències d'autoavaluació o de reflexió i el seu comentari
- Evidències de l'evolució realitzada
- Objectius futurs
- Agraïments

El procés d'elaboració d'un portafolis depèn de les orientacions que proposi la persona o la institució que el sol·liciti. En general, s'inicia amb la recopilació del material que serveix com a evidència de l'aprenentatge; en un primer moment, es recomana de no ser gaire selectiu. Un altre pas clau és decidir com s'organitzen els comentaris o les reflexions —per temes, per etapes, per importància, etc. Es recomana anar discutint el portafolis mentre s'elabora amb algun company o tutor. La selecció final del material ha de ser bastant discriminativa, ja que importa més la qualitat que la quantitat.

La persona que ha d'avaluar un portafolis ha de tenir en compte que, tant l'elaboració com l'avaluació del portafolis, consumeixen força temps. Un criteri possible per a l'avaluació és el personalitzat, és a dir, comparar els resultats amb els treballs fets anteriorment per la mateixa persona; per a això cal que tot el material inclòs a la carpeta estigui datat. Algunes recomanacions per a la persona que faci l'avaluació són:

- Especificar o negociar clarament els resultats d'aprenentatge que s'esperen.
- Proposar un format general de portafolis.
- Especificar o negociar la naturalesa de les evidències que s'espera trobar al portafolis.
- Especificar o negociar l'extensió de les evidències.
- Preparar una llista de comprovació o escala d'estimació per avaluar el portafolis.
- Utilitzar notes adhesives —en paper o digitals— per marcar aquelles parts del portafolis que cal repassar.
- Oferir possibilitats d'autoavaluació durant el procés d'elaboració del portafolis.

La utilització de portafolis no està lliure de conflictes. Alguns estudiants opinen que és una tasca massa dura i que és el professorat qui ha de dir què sap l'alumne. De vegades es crea una tensió entre el que és oportú d'utilitzar educativament per avaluar i el que es requereix des de fora de l'aula o del sistema educatiu. La resposta a aquests conflictes ens redirigeix de nou al concepte d'educació: no es tracta simplement de canviar «la manera de fer», sinó de transformar la pràctica de forma reflexiva. El paper del professorat, llavors, consisteix sobretot a facilitar i orientar el procés reflexiu, autoavaluat, formatiu i educatiu que comporta l'elaboració d'un portafolis.

Tal com s'ha comentat anteriorment sobre les proves alternatives de rendiment, els portafolis tenen una utilitat —entesa com a aplicabilitat o validesa— raonablement alta. En canvi, la consistència (o fiabilitat) d'aquest instrument és una àrea encara per investigar. Baume i Yorke (2002) descriuen un estudi sobre la consistència dels avaluadors de portafolis docents —portafolis que elabora el professorat per al seu propi desenvolupament professional— i conclouen que cal formar els avaluadors per construir una comprensió compartida del que significa un portafolis i uns criteris d'avaluació. De totes maneres, el valor formatiu del portafolis fa que la seva consistència no sigui una qüestió tan important.

5 Bibliografia per saber-ne més

- **Proves d'assaig** (ESPÍN i RODRÍGUEZ LAJO, 1996, p. 67-77; MATEO i MARTÍNEZ-OLMO, 2008b, p. 100-102).
 - Problemes amb les proves d'assaig (CURZON, 2002, p. 393-394).
- **Proves objectives** (BLANCO, 1994, p. 108-112; COHEN et al., 2000, p. 317-336; DEL RINCÓN et al., 1995, p. 141-176; ESPÍN i RODRÍGUEZ LAJO, 1996, p. 79-112; MATEO i MARTÍNEZ-OLMO, 2008b, p. 102-117; RENOM, 1992; SANS, 2004, p. 32-43, 2008; TENBRINK, 1997, p. 313-374).
- **Proves alternatives i portafolis** (MATEO, 2000, p. 82-85; MATEO i MARTÍNEZ-OLMO, 2005, 2008a, p. 209-263).
 - Rúbriques (LÓPEZ, 2002; STEVENS i LEVI, 2005; SUSKIE, 2009, p. 137-154).

6 Referències

- ARTER, Judith A. (1995). *Portfolios for Assessment and Instruction*. ERIC Digest, ED388890.
- BAUME, David, YORKE, Mantz. (2002). «The Reliability of Assessment by Portfolio on a Course to Develop and Accredite Teachers in Higher Education». *Studies in Higher Education*, 27(1), 7–25. <<https://doi.org/10.1080/03075070120099340>>.
- BLANCO, Francisco. (1994). *La evaluación en la educación secundaria*. Amarú.
- COHEN, Louis, MANION, Lawrence, MORRISON, Keith. (2000). *Research methods in education*. 5a ed. RoutledgeFalmer.
- COHEN, Louis, MANION, Lawrence, MORRISON, Keith. (2007). *Research methods in education*. 6a ed. Routledge.
- CURZON, Leslie Basil. (2002). *Teaching in Further Education. An Outline of Principles and practice*. 5a ed. Continuum.
- DEL RINCÓN, Delio, ARNAL, Justo, LATORRE, Antonio, SANS, Antoni. (1995). *Técnicas de investigación en ciencias sociales*. Dykinson.
- ESPÍN, Julia Victoria, RODRÍGUEZ LAJO, Mercedes. (1996). *L'avaluació dels aprenentatges a la universitat*. Publicacions de la Universitat de Barcelona.
- LÓPEZ, Juan Carlos. (2002). *Matriz de Valoración (Rúbricas - Rubrics en inglés)*. Eduteka. <<http://eduteka.icesi.edu.co/articulos/MatrizValoracion>>.
- MATEO, Joan. (2000). *La evaluación educativa, su práctica y otras metáforas*. ICE y Horsori.
- MATEO, Joan, MARTÍNEZ-OLMO, Francesc. (2005). *L'avaluació alternativa dels aprenentatges* (vol. 3). ICE. Universitat de Barcelona. <<https://doi.org/10.1344/101.8488795858>>.
- MATEO, Joan, MARTÍNEZ-OLMO, Francesc. (2008a). *La evaluación alternativa de los aprendizajes*. ICE-UB y Octaedro. <<https://octaedro.com/libro/la-evaluacion-alternativa-de-los-aprendizajes/>>.
- MATEO, Joan, MARTÍNEZ-OLMO, Francesc. (2008b). *Medición y evaluación educativa*. La Muralla.
- MEYER, Carol. A. (1992). «What's the Difference Between Authentic and Performance Assessment?». *Educational Leadership*, 49(8), 39-40. <http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_199205_meyer.pdf>.

- PAPPAS, Peter. (2010). *The Reflective Teacher: A Taxonomy of Reflection (Part 3)*. <<https://peterpappas.com/2010/01/reflective-teacher-taxonomy-reflection.html>>.
- PETT, Joel. (1990). «What is Authentic Evaluation? Common Questions and Answers». *FairTest Examiner*, 4, 8–9.
- RENOM, Jordi. (1992). *Diseño de tests*. Idea Investigación y Desarrollo.
- SANS, Antoni. (2004). *L'avaluació dels aprenentatges: construcció d'instruments*. Universitat de Barcelona. Institut de Ciències de l'Educació. <<http://hdl.handle.net/2445/8447>>.
- SANS, Antoni. (2008). *La evaluación de los aprendizajes: construcción de instrumentos*. ICE-UB y Octaedro. <<http://hdl.handle.net/2445/144919>>.
- STEVENS, Danelle D., LEVI, Antonia J. (2005). *Introduction to Rubrics*. Stylus.
- SUNSTEIN, Bonnie S., LOVELL, Jonathan H. (Eds.). (2000). *The portfolio standard. How students can show us what they know and are able to do*. Heinemann.
- SUSKIE, Linda. (2009). *Assessing Student Learning*. 2a ed. Jossey-Bass.
- TENBRINK, Terry. D. (1997). *Evaluación. guía práctica para profesores*. 4a ed. Narcea.
- WOLF, Dennie P., REARDON, Sean F. (1996). *Access to excellence through new forms of student assessment*. En Joan B. BARON, Dennie. P. WOLF (Eds.), *Performance-based student assessment: Challenges and possibilities. Ninty-fifth yearbook of the National Society for the Study of Education* (pp. 52–83). University of Chicago Press.